Package 'rplec'

January 23, 2025

Title Placental Epigenetic Clock to Estimate Aging by DNA Methylation

Version 0.1.3

Description Placental epigenetic clock to estimate aging based on gestational age using DNA methylation levels, so called placental epigenetic clock (PlEC). We developed a PlEC for the 2024 Placental Clock DREAM Challenge (<https://www.synapse.org/Synapse:syn59520082/wiki/628063>). Our PlEC achieved the top performance based on an independent test set. PlEC can be used to identify accelerated/decelerated aging of placenta for understanding placental dysfunction-related conditions, e.g., great obstetrical syndromes including preeclampsia, fetal growth restriction, preterm labor, preterm premature rupture of the membranes, late spontaneous abortion, and placental abruption. Detailed methodologies and examples are documented in our vignette, available at <https://herdiantrisufriyana.github.io/rplec/doc/placental_aging_analysis.html>.

Depends R (>= 4.4)

Imports doParallel,
dplyr,
foreach,
ggplot2,
parallel,
pbapply,
purrr,
RPMM,
stats,
stringr,
tibble,
tidyr,
utils

Suggests tidyverse,
knitr,
testthat (>= 3.0.0)

VignetteBuilder knitr

License MIT + file LICENSE

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.2

LazyData true

LazyDataCompression xz

Config/testthat/edition 3

Contents

---

aging | *Residual DNA-methylation-based (DNAm) gestational age*

Description

A data frame of DNAm gestational age for both case and control groups.

Usage

```
aging
```

Format

A data frame vector with 10 rows and 1 column:

output A numeric value of residual DNAm gestational age (weeks).

Source

Derived from the 2024 Placental Clock DREAM Challenge.

---

beta_v_indices *Indices for random beta values*

Description

A list of indices for rand.idx in bmiq_norm_450k function. The input probes must be already filtered and ordered the same way to the that when we developed our placental epigenetic clock. Run data(probe_info_450k) and find the required probes in prob_info_450k$probeID.

Usage

```
beta_v_indices
```

Format

A list of 2 elements where each has a length of nfit of 10000:

beta1.v An integer indicating the selected indices.

beta2.v An integer indicating the selected indices.

Source

Derived from ChAMP R package.

---

bmiq_norm_450k *Normalize DNA methylation values*

Description

This function normalize DNA methylation values from 450k probes. The probes are filtered and ordered the same way to the input when we developed our placental epigenetic clock.

Usage

```
bmiq_norm_450k(beta, cores = 1, verbose = FALSE)
```

Arguments

beta	A data frame of beta values where each column represents a sample and each row represent a probe. The rows must be named according to the probe IDs. They include all the required probes. Run data(probe_info_450k) and find the required probes in prob_info_450k$probeID. Meanwhile, the columns must be named according to the sample IDs.
cores	An integer indicating the number of threads.
verbose	A logical scalar indicating whether to show a progress bar.

Value

A data frame of normalized beta values.

Examples

```
beta_values_case <- download_beta_values_case()
norm_beta_values_case <- bmiq_norm_450k(beta_values_case)
```

---

download_beta_values_case

Access Beta values at individual CpG sites for case group

Description

Downloads and loads beta values for the case group. Data contains beta values for 5 samples and 452,453 probes.

Usage

```
download_beta_values_case()
```

Value

A data frame with 452,453 rows and 5 columns:

GSM1931565 Beta values for sample GSM1931565.

GSM5114811 Beta values for sample GSM5114811.

GSM2589558 Beta values for sample GSM2589558.

GSM1842848 Beta values for sample GSM1842848.

GSM1843045 Beta values for sample GSM1843045.

Source

Derived from the 2024 Placental Clock DREAM Challenge.

Examples

```
beta_values_case <- download_beta_values_case()
head(beta_values_case)
```

```
download_beta_values_control
```
Access Beta values at individual CpG sites for control group

Description

Downloads and loads beta values for the control group. Data contains beta values for 5 samples and 452,453 probes.

Usage

```
download_beta_values_control()
```

Value

A data frame with 452,453 rows and 5 columns:

GSM7115144 Beta values for sample GSM7115144.

GSM1702248 Beta values for sample GSM1702248.

GSM3179749 Beta values for sample GSM3179749.

GSM4281756 Beta values for sample GSM4281756.

GSM5210472 Beta values for sample GSM5210472.

Source

Derived from the 2024 Placental Clock DREAM Challenge.

Examples

```
beta_values_control <- download_beta_values_control()
head(beta_values_control)
```

---

ga *Gestational age for both case and control groups*

Description

A data frame of gestational age for 10 samples.

Usage

```
ga
```

Format

A data frame vector with 10 rows and 1 column:

GA A numeric value of gestational age (weeks' gestation).

Source

Derived from the 2024 Placental Clock DREAM Challenge.

---

ipla   *Identify placental aging*

Description

This function identifies placental aging based on the case-control aging difference. Placental aging is defined as the residual DNA-methylation-based (DNAm) gestational ages (GA). Only GA from 5 to 44 weeks' gestation are shown in the placental aging plot.

Usage

```
ipla(
  aging,
  ga,
  phenotype,
  case = "Case",
  control = "Control",
  method = NULL,
  from = NULL,
  to = NULL
)
```

Arguments

aging	A data frame of residual DNA-methylation-based GA. This data frame must be the output of `plec` function with argument `type="residual"`. The rows must be named according to the sample IDs.
ga	A data frame of GA. There is only one column, i.e., GA, as shown in ga. Use `data(ga)` to load this data frame. The rows must be named according to the sample IDs.
phenotype	A data frame of phenotype (optional. There is only one column, i.e., `phenotype`, as shown in `phenotype`. Use `data(phenotype)` to load this data frame. The rows must be named according to the sample IDs.
case	A character of of case name in phenotype (default="Case").
control	A character of of case name in phenotype (default="Control").
method	A character of of the method of statistical test (optional), i.e., "Mann-Whitney U" or "Permutation".

from	An integer from 5 to 44 indicating minimum GA (weeks) to be included in the statistical test. If it is undefined, the minimum GA in either case or control is applied.
to	An integer from 5 to 44 indicating maximum GA (weeks) to be included in the statistical test. If it is undefined, the maximum GA in either case or control is applied.

Value

An ggplot object consisting the aging plot without or with statistical test results.

Examples

```
# Prepare data
data(aging)
data(ga)
data(phenotype)

# Identify placental aging
set.seed(1)
ipla(aging, ga, phenotype)

## Conduct statistical test
set.seed(1)
ipla(aging, ga, phenotype, method = "Mann-Whitney U")

## Conduct statistical test for a specific range of GA
set.seed(1)
ipla(aging, ga, phenotype, method = "Mann-Whitney U", from = 5, to = 20)
```

---

phenotype	*Phenotype for both case and control groups*

Description

A data frame of phenotype for 10 samples.

Usage

```
phenotype
```

Format

A data frame vector with 10 rows and 1 column:

phenotype A character value of phenotype (case/control).

Source

Derived from the 2024 Placental Clock DREAM Challenge.

---

plec   *Estimate DNA-methylation-based gestational age*

Description

This function estimate gestational age (GA) using BMIQ-normalized beta values. The estimated GA is a sum of normal and residual GAs. The latter is a sum of condition- and trimester-specific, residual GAs.

Usage

```
plec(norm_beta, type = "stack", verbose = FALSE)
```

Arguments

norm_beta   A data frame of normalized beta values where each column represents a sample and each row represent a probe. This data frame must be the output of [bmiq_norm_450k](#) function. The rows must be named according to the probe IDs. Meanwhile, the columns must be named according to the sample IDs.

type   An character indicating the type of outputs which are primarily: (1) "stack" (default) for the estimated GA; (2) "normal" for the estimated normal GA; (3) "residual" for the estimated residual GA; (4) "condition" for the condition-specific, estimated residual GA; and (5) "trimester" for the trimester-specific, estimated residual GA. In addition, a user can obtain the output of a single sub-model using the column name (except predictor) in plec_int_coef. Use data(plec_int_coef) to load this data frame.

verbose   A logical scalar indicating whether to show a progress bar.

Value

A data frame of the estimated GA.

Examples

```
beta_values_case <- download_beta_values_case()
norm_beta_values_case <- bmiq_norm_450k(beta_values_case)
dnam_ga_case <- plec(norm_beta_values_case)
```

---

plec_int_coef	*Intercept and coefficients of placental epigenetic clock*

Description

A data frame of intercept and coefficients for all submodels in our placental epigenetic_clock.

Usage

```
plec_int_coef
```

Format

A data frame with 10,447 rows and 32 columns:

predictor A character value of predictor name.

ga_est Estimate normal GA.

fgr_pred Predict FGR.

pe_pred Predict PE.

pe_onset_pred Predict PE onset.

preterm_pred Predict preterm delivery.

anencephaly_pred Predict anencephaly.

spina_bifida_pred Predict spina bifida.

gdm_pred Predict GDM.

diandric_triploid_pred Predict diandric triploid.

miscarriage_pred Predict miscarriage.

lga_pred Predict LGA.

subfertility_pred Predict subfertility.

hellp_pred Predict HELLP.

chorioamnionitis_pred Predict chorioamnionitis.

ga_res_conds_fgr_est Estimate GA in FGR.

ga_res_conds_pe_est Estimate GA in PE.

ga_res_conds_pe_onset_est Estimate GA in EOPE.

ga_res_conds_preterm_est Estimate GA in preterm delivery.

ga_res_conds_anencephaly_est Estimate GA in anencephaly.

ga_res_conds_spina_bifida_est Estimate GA in spina bifida.

ga_res_conds_gdm_est Estimate GA in GDM.

ga_res_conds_diandric_triploid_est Estimate GA in diandric triploid.

ga_res_conds_miscarriage_est Estimate GA in miscarriage.

ga_res_conds_lga_est Estimate GA in LGA.

ga_res_conds_subfertility_est Estimate GA in subfertility.

ga_res_conds_hellp_est Estimate GA in HELLP.

ga_res_conds_chorioamnionitis_est Estimate GA in chorioamnionitis.

ga_res_conds_pred_est Estimate residual GA.

ga_res_comb_pr_est Estimate residual GA for preterm.

ga_res_comb_tb_est Estimate residual GA for term before the date.

ga_res_comb_ta_est Estimate residual GA for term after the date.

Source

Derived from the 2024 Placental Clock DREAM Challenge.

---

plec_scaler_mean   *Mean values of the scalers of placental epigenetic clock*

Description

A data frame of mean values of the scalers for all submodels in our placental epigenetic_clock.

Usage

```
plec_scaler_mean
```

Format

A data frame with 10,446 rows and 32 columns:

predictor A character value of predictor name.

ga_est Estimate normal GA.

fgr_pred Predict FGR.

pe_pred Predict PE.

pe_onset_pred Predict PE onset.

preterm_pred Predict preterm delivery.

anencephaly_pred Predict anencephaly.

spina_bifida_pred Predict spina bifida.

gdm_pred Predict GDM.

diandric_triploid_pred Predict diandric triploid.

miscarriage_pred Predict miscarriage.

lga_pred Predict LGA.

subfertility_pred Predict subfertility.

hellp_pred Predict HELLP.

chorioamnionitis_pred Predict chorioamnionitis.

ga_res_conds_fgr_est Estimate GA in FGR.

ga_res_conds_pe_est Estimate GA in PE.

ga_res_conds_pe_onset_est Estimate GA in EOPE.

ga_res_conds_preterm_est Estimate GA in preterm delivery.

ga_res_conds_anencephaly_est Estimate GA in anencephaly.

ga_res_conds_spina_bifida_est Estimate GA in spina bifida.

ga_res_conds_gdm_est Estimate GA in GDM.

ga_res_conds_diandric_triploid_est Estimate GA in diandric triploid.

ga_res_conds_miscarriage_est Estimate GA in miscarriage.

ga_res_conds_lga_est Estimate GA in LGA.

ga_res_conds_subfertility_est Estimate GA in subfertility.

ga_res_conds_hellp_est Estimate GA in HELLP.

ga_res_conds_chorioamnionitis_est Estimate GA in chorioamnionitis.

ga_res_conds_pred_est Estimate residual GA.

ga_res_comb_pr_est Estimate residual GA for preterm.

ga_res_comb_tb_est Estimate residual GA for term before the date.

ga_res_comb_ta_est Estimate residual GA for term after the date.

Source

Derived from the 2024 Placental Clock DREAM Challenge.

| plec_scaler_scale | *Scale values of the scalers of placental epigenetic clock* |

Description

A data frame of scale values of the scalers for all submodels in our placental epigenetic_clock.

Usage

```
plec_scaler_scale
```

Format

A data frame with 10,446 rows and 32 columns:

predictor A character value of predictor name.

ga_est Estimate normal GA.

fgr_pred Predict FGR.

pe_pred Predict PE.

pe_onset_pred Predict PE onset.

preterm_pred Predict preterm delivery.

anencephaly_pred Predict anencephaly.

spina_bifida_pred Predict spina bifida.

gdm_pred Predict GDM.

diandric_triploid_pred Predict diandric triploid.

miscarriage_pred Predict miscarriage.

lga_pred Predict LGA.

subfertility_pred Predict subfertility.

hellp_pred Predict HELLP.

chorioamnionitis_pred Predict chorioamnionitis.

ga_res_conds_fgr_est Estimate GA in FGR.

ga_res_conds_pe_est Estimate GA in PE.

ga_res_conds_pe_onset_est Estimate GA in EOPE.

ga_res_conds_preterm_est Estimate GA in preterm delivery.

ga_res_conds_anencephaly_est Estimate GA in anencephaly.

ga_res_conds_spina_bifida_est Estimate GA in spina bifida.

ga_res_conds_gdm_est Estimate GA in GDM.

ga_res_conds_diandric_triploid_est Estimate GA in diandric triploid.

ga_res_conds_miscarriage_est Estimate GA in miscarriage.

ga_res_conds_lga_est Estimate GA in LGA.

ga_res_conds_subfertility_est Estimate GA in subfertility.

ga_res_conds_hellp_est Estimate GA in HELLP.

ga_res_conds_chorioamnionitis_est Estimate GA in chorioamnionitis.

ga_res_conds_pred_est Estimate residual GA.

ga_res_comb_pr_est Estimate residual GA for preterm.

ga_res_comb_tb_est Estimate residual GA for term before the date.

ga_res_comb_ta_est Estimate residual GA for term after the date.

Source

Derived from the 2024 Placental Clock DREAM Challenge.

---

probe_info_450k *Probe info for 450K*

Description

A list of 450K probe information for `bmiq_norm_450k` function. The probes are already filtered and ordered the same way to the input when we developed our placental epigenetic clock.

Usage

```
probe_info_450k
```

Format

A list of 2 elements where each has a length of 346407:

Design An integer indicating design type 1 or 2.

probeID A character for each probe identifier.

Source

Derived from `ChAMP` R package.

---

qc *Perform quality control*

Description

This function evaluates the precision of DNA-methylation-based (DNAm) gestational age (GA) based on calibration, root mean square error (RMSE), mean absolute error (MAE), and Pearson's correlation coefficient (r). The sample identifiers (IDs) are automatically matched among the DNAm-GA, GA, and phenotype (optional). Only GA from 5 to 44 weeks' gestation are shown in the calibration plot.

Usage

```
qc(dnam_ga, ga, phenotype = NULL)
```

Arguments

dnam_ga	A data frame of DNA-methylation-based GA. This data frame must be the output of `plec` function. The rows must be named according to the sample IDs.
ga	A data frame of GA. There is only one column, i.e., GA, as shown in ga. Use `data(ga)` to load this data frame. The rows must be named according to the sample IDs.
phenotype	A data frame of phenotype (optional). There is only one column, i.e., phenotype, as shown in phenotype. Use `data(phenotype)` to load this data frame. The rows must be named according to the sample IDs.

Value

A ggplot object of calibration plot with RMSE, MAE, and r.

Examples

```
beta_values_case <- download_beta_values_case()
norm_beta_values_case <- bmiq_norm_450k(beta_values_case)
dnam_ga_case <- plec(norm_beta_values_case)

data(ga)
ga_case <- ga[phenotype$phenotype == "Case", , drop = FALSE]
set.seed(1)
qc(dnam_ga_case, ga_case)
```

Index